Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2007, Article ID 70351, 9 pages
doi:10.1155/2007/70351

Research Article

Recognition of Planar Objects Using Multiresolution Analysis

Nazlı Güney and Ayşın Ertüzün

Department of Electrical and Electronics Engineering, Boğaziçi University, 34342 Bebek, Istanbul, Turkey

Received 29 August 2005; Revised 29 May 2006; Accepted 16 July 2006

Recommended by Antonio Ortega

By using affine-invariant shape descriptors, it is possible to recognize an unknown planar object from an image taken from an arbitrary view when standard view images of candidate objects exist in a database. In a previous study, an affine-invariant function calculated from the wavelet coefficients of the object boundary has been proposed. In this work, the invariant is constructed from the multiwavelet and (multi)scaling function coefficients of the boundary. Multiwavelets are known to have superior performance compared to scalar wavelets in many areas of signal processing due to their simultaneous orthogonality, symmetry, and short support properties. Going from scalar wavelets to multiwavelets is challenging due to the increased dimensionality of multiwavelets. This increased dimensionality is exploited to construct invariants with better performance when the multiwavelet "detail" coefficients are available. However, with (multi)scaling function coefficients, which are more stable in the presence of noise, scalar wavelets cannot be defeated.

1. INTRODUCTION

Object recognition is one of the most difficult problems in computer vision. However, if the problem definition includes only planar objects, which are to be viewed from arbitrary directions, it is possible to design recognition systems that have satisfactory performances. When the depth of an object along the line of sight is small compared to the viewing distance of the camera, as its images are produced from different viewpoints, it seems to be going through an affine transformation. Thus, for recognition of planar objects, it suffices to find suitable affine invariants. These invariants are shape descriptors that remain unchanged even when the viewing point of the camera changes. Therefore, it may be said that object recognition is a search for invariants [1].

Recognition techniques are classified according to how the shape descriptors are calculated from the images of objects. One such classification is based on whether the boundary or the region of the object is required. Region-based techniques take into account the whole region in the image corresponding to the object, whereas boundary-based techniques analyze the object boundary. Analyzing only the boundary is advantageous compared to the region-based techniques in terms of computational complexity, since the amount of data to be processed substantially diminishes. Yet another classification to discriminate between the shape descriptors is whether they are local or global. Local techniques, which usually resort to higher-order derivatives, are very much affected by the presence of noise [1]. Global techniques, on the other hand, consider the whole data when calculating the shape descriptors, and hence suffer from occlusion of the object. Fourier descriptors in [2] and the wavelet-transform-based methods [3–5], which are the subject of this paper, are examples of boundary-based global techniques, since computation of the transform coefficients requires all of the boundary coordinates.

Among the wavelet-transform-based techniques, the work by Khalil and Bayoumi [4] deserves further attention. When calculating the affine-invariant shape descriptor, they have used the biggest number of wavelet scales. The affine-invariant function proposed in [4] uses 7–12 wavelet scales. Different boundaries may have similar wavelet coefficients at a particular scale, but not at all scales [4]. Thus, with more scales used, the more accurate the representation of the boundary becomes. A database of 20 airplane objects is used to test the recognition performance of the affine-invariant function. Uniformly distributed noise at 50 dB and 20 dB signal-to-noise ratios (SNR) is added to the randomly affine-transformed object boundaries [4, 6]. However, only one realization of noise and one object view are not sufficient to assess the performance of the invariant function proposed. Besides, only one type of wavelet, the one in [7], has been used in the simulations accompanied with an inadequate analysis as to which wavelet scales should be chosen.

Multiwavelets, which are generalizations of wavelets, have shown superior performance compared to wavelets in such areas as image compression [8] and image denoising [9, 10]. These application areas are related to object recognition, for they, too, require compact and accurate representations. Thus, the affine-invariant function should also benefit from using multiwavelet coefficients instead of wavelet coefficients. Since coefficients at different scales are multiplied together when calculating the invariant function, the undecimated (redundant) multiwavelet transform, which is also translation invariant, is employed. In [11], where planar shapes are represented with the orthogonal multiwavelet transform coefficients, the multiwavelets are shown to be more promising than scalar wavelets in terms of accuracy of representation.

In this work, the affine-invariant function in [4] is constructed from (multi)wavelet and (multi)scaling function coefficients of the object boundary. Extensive simulations are made with three object databases and hundred views for each of the objects. Four wavelets, two combined sets of wavelets, and six multiwavelets are tested. The approximation properties of the (multi)wavelets are shown to be the most significant criterion when choosing either the transform coefficient scales or the type of (multi)wavelet. Moreover, whether the objects are smooth or contain detail affect the performance of the invariant function.

The rest of the paper is organized as follows. Section 2 reviews the orthogonal multiwavelet transform and explains the procedure for calculating the redundant multiwavelet transform. The affine-invariant function in [4] is introduced in Section 3, where the differences between the invariants using (multi)wavelet and (multi)scaling function coefficients are outlined. Experimental results and a theoretical analysis based on the approximation properties of (multi)wavelets are in Section 4. Finally, conclusions are made in Section 5.

2. MULTIWAVELET TRANSFORM

2.1. Orthogonal multiwavelet transform

Multiwavelets have been introduced as an extension to scalar wavelets and are defined by a set of wavelets instead of a single wavelet [12]. The theory is, again, based on the idea of multiresolution analysis [13]. The standard multiresolution, which has one scaling function, $\phi(t)$, has the following properties.

(i) The translates $\phi(t-k)$ are linearly independent and produce a basis for the subspace V_0.

(ii) The dilates $\phi(2^j t-k)$ generate subspaces V_j such that

$$\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots \subset V_j \subset \cdots,$$
$$\overline{\bigcup_{j=-\infty}^{\infty} V_j} = L^2(\mathbf{R}), \qquad \bigcap_{j=-\infty}^{\infty} V_j = \{0\}, \tag{1}$$

where $L^2(\mathbf{R})$ is the vector space of measurable, square integrable one-dimensional (1D) functions.

(iii) The integer translates of the wavelet $\psi(t-k)$ produce a basis for the "detail" subspace W_0 to give V_1,

$$V_1 = V_0 \oplus W_0, \tag{2}$$

where $V_0 \perp W_0$.

For multiwavelets, the subspace V_0 is spanned by translates of R scaling functions. The resulting multiscaling function is defined as a column vector, where each row corresponds to a scaling function: $\mathbf{\Phi}(t) = [\phi_1(t), \ldots, \phi_R(t)]^T$. The related multiwavelet is $\mathbf{\Psi}(t) = [\psi_1(t), \ldots, \psi_R(t)]^T$. Multiwavelets have $R \geq 2$, and with $R = 1$, scalar wavelets are obtained. The relationship between multiscaling functions at adjacent scales is described with the matrix refinement equation [14]

$$\mathbf{\Phi}(t) = \sqrt{2} \sum_k \mathbf{H}(k) \mathbf{\Phi}(2t-k). \tag{3}$$

Similarly, the multiwavelet is expressed as a weighted sum of the multiscaling functions at the next finer scale,

$$\mathbf{\Psi}(t) = \sqrt{2} \sum_k \mathbf{G}(k) \mathbf{\Phi}(2t-k). \tag{4}$$

In (3) and (4), the low-pass filter coefficients $\mathbf{H}(k)$ and the high-pass filter coefficients $\mathbf{G}(k)$ are $R \times R$ matrices.

If $f(t) \in V_{J+1}$, it can be written as a linear combination of multiscaling functions and multiwavelets with

$$f(t) = \sum_k \mathbf{C}_{j_0}^T(k) \mathbf{\Phi}_{j_0,k}(t) + \sum_{j=j_0}^{J} \sum_k \mathbf{D}_j^T(k) \mathbf{\Psi}_{j,k}(t), \tag{5}$$

where $\mathbf{C}_j$ and $\mathbf{D}_j$ represent the multiscaling function and multiwavelet coefficients at scale j, respectively, j_0 denotes the coarsest scale and

$$\mathbf{\Phi}_{j_0,k}(t) = 2^{j_0/2} \mathbf{\Phi}\left(2^{j_0} t - k\right), \qquad \mathbf{\Psi}_{j,k}(t) = 2^{j/2} \mathbf{\Psi}\left(2^j t - k\right) \tag{6}$$

are $\mathbf{\Phi}(t)$ and $\mathbf{\Psi}(t)$ shifted in time and then scaled in amplitude and time, respectively. Since multiwavelets and multiscaling functions are orthogonal, the multiscaling function ($\mathbf{C}$) and multiwavelet coefficients ($\mathbf{D}$) at a coarser scale can be calculated from the multiscaling function coefficients at a finer scale,

$$\mathbf{C}_{j-1}(k) = \sqrt{2} \sum_m \mathbf{H}(m-2k) \mathbf{C}_j(m),$$
$$\mathbf{D}_{j-1}(k) = \sqrt{2} \sum_m \mathbf{G}(m-2k) \mathbf{C}_j(m). \tag{7}$$

These are the analysis equations that can be implemented with a filterbank consisting of low- and high-pass filters followed with downsamplers. This is demonstrated in Figure 1. Each single scaling function of the multiscaling function has a scalar coefficient associated with it. Hence, $\mathbf{C}$ is a column vector of dimension R. In order to start the filterbank, initial estimates of the multiscaling function coefficients at the finest (highest) scale have to be obtained from the samples of the signal $f(t)$. The signal samples are, thus, preprocessed

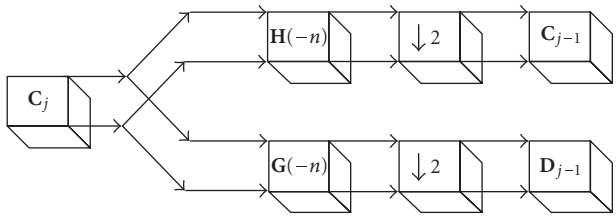

Figure 1: The analysis filterbank for orthogonal multiwavelet transform.

(prefiltered) to produce reasonable values for the coefficients of the multiscaling function at the finest scale [14]. A number of preprocessing techniques have been proposed for this purpose. Repeated row (RR) and approximation (AP) preprocessings are the two most widely used techniques [15]. In RR preprocessing, the rows of the input vector to the filterbank are obtained by scaling the first row consisting of the signal samples. This preprocessing increases the total number of samples leading to an oversampling of the original signal. AP preprocessing, which is based on the approximation properties of continuous-time wavelets, on the other hand, yields a critically sampled representation.

2.2. Redundant (undecimated) multiwavelet transform

The procedure for calculating the redundant multiwavelet transform is based on the work of Mallat in [16] for scalar wavelets. If $R = 2$ and RR (AP) preprocessing is employed on a 1D signal with length N, the preprocessed signal is $2 \times N$ ($2 \times N/2$). $\mathbf{H}(k)$ are 2×2 matrices and $\mathbf{C}_j(k)$ are 2×1 vectors,

$$\mathbf{H}(k) = \begin{bmatrix} h_1(k) & h_2(k) \\ h_3(k) & h_4(k) \end{bmatrix}, \qquad \mathbf{C}_j(k) = \begin{bmatrix} c_{j,1}(k) \\ c_{j,2}(k) \end{bmatrix}. \tag{8}$$

Regarding each element of a matrix filter coefficient as the coefficient of a scalar filter, $h_1(-k)$ and $h_3(-k)$ filter the first, and $h_2(-k)$ and $h_4(-k)$ filter the second row of multiscaling function coefficients,

$$\begin{aligned} c_{j-1,1}(k) &= c_{j,1}(k) * h_1(-k) + c_{j,2}(k) * h_2(-k), \\ c_{j-1,2}(k) &= c_{j,1}(k) * h_3(-k) + c_{j,2}(k) * h_4(-k), \end{aligned} \tag{9}$$

where $*$ denotes convolution. In this form, it is apparent that both rows of multiwavelet and multiscaling function coefficients at a coarser scale depend on both rows of multiscaling function coefficients at a finer scale. We have obtained the redundant multiwavelet transform by avoiding downsampling and padding each of the scalar filters with zeros for upsampling. This has the same effect as padding each of the matrix filter coefficients with zero matrices of size 2×2. Consequently, the number of redundant multiwavelet coefficients at each scale is identical (i.e., $2 \times N$).

3. MULTIRESOLUTION ANALYSIS OF THE OBJECT BOUNDARY

When a planar object is to be recognized from its image, the boundary of the object, which is modeled with a 2D curve, is analyzed. Consider a situation where reference images of the objects to be recognized are kept in a database, and a test image taken from a different view of one of the objects is also present. The goal is to find to which object this test image belongs. Each point $(x(t), y(t))$ on the boundary curve in the reference image has been mapped to a point $(\tilde{x}(t), \tilde{y}(t))$ on the curve in the test image,

$$\begin{aligned} \tilde{x}(t) &= a_0 + a_1 x(t) + a_2 y(t), \\ \tilde{y}(t) &= b_0 + b_1 x(t) + b_2 y(t). \end{aligned} \tag{10}$$

Formulas (10) are combined as

$$\tilde{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}, \tag{11}$$

where

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix}. \tag{12}$$

$\mathbf{A}$ in (12) is a nonsingular square matrix representing rotation, scaling, and skewing in the affine transformation, and vector $\mathbf{b}$ represents translation.

3.1. The affine-invariant wavelet function

The wavelet coefficients at scale j of the boundary curve in the test image are related to those of the boundary curve in the reference image by an equation similar to (11):

$$W_j\tilde{\mathbf{x}} = \mathbf{A}W_j\mathbf{x}, \tag{13}$$

where W_j denotes the wavelet "detail" coefficients at scale j. More clearly,

$$\begin{aligned} W_j\tilde{x}(t) &= a_1 W_j x(t) + a_2 W_j y(t), \\ W_j\tilde{y}(t) &= b_1 W_j x(t) + b_2 W_j y(t), \end{aligned} \tag{14}$$

with $W_j a_0 = W_j b_0 = 0$ due to high-pass filtering. An affine-invariant function is defined in [4] with two scales as

$$f_{i,j}(t) = W_i x(t) W_j y(t) - W_i y(t) W_j x(t), \quad i \neq j. \tag{15}$$

This is a relative invariant, where different affine transformations of the boundary produce scaled versions of $f_{i,j}(t)$, since it is given by

$$\begin{aligned} \det\{[W_i\tilde{\mathbf{x}} \;\; W_j\tilde{\mathbf{x}}]\} &= \det\{\mathbf{A}[W_i\mathbf{x} \;\; W_j\mathbf{x}]\}, \\ \tilde{f}_{i,j}(t) &= \det(\mathbf{A}) f_{i,j}(t), \end{aligned} \tag{16}$$

where $\det(\cdot)$ is the determinant.

An affine-invariant function with six wavelet scales is proposed in [4] by introducing a wavelet-based conic equation using three wavelet scales. The shape descriptor is the invariant of two wavelet-based conics with parametrized coefficients, where the conics are defined for the scales $\{a, b, c\}$ and $\{d, e, f\}$. Different wavelet-based conics are represented with the two sets of scales. When the coefficients of the conics are solved for, the function $\eta_{a,b,c,d,e,f}(t)$ calculated from

wavelet scales $\{a, b, c, d, e, f\}$ is obtained [4]:

$$\begin{aligned}\eta_{a,b,c,d,e,f}(t) &= \begin{vmatrix} 1 & 2W_a x W_a y & W_a^2 y \\ 1 & 2W_b x W_b y & W_b^2 y \\ 1 & 2W_c x W_c y & W_c^2 y \end{vmatrix} \begin{vmatrix} W_d^2 x & 2W_d x W_d y & 1 \\ W_e^2 x & 2W_e x W_e y & 1 \\ W_f^2 x & 2W_f x W_f y & 1 \end{vmatrix} \\ &+ \begin{vmatrix} W_a^2 x & 2W_a x W_a y & 1 \\ W_b^2 x & 2W_b x W_b y & 1 \\ W_c^2 x & 2W_c x W_c y & 1 \end{vmatrix} \begin{vmatrix} 1 & 2W_d x W_d y & W_d^2 y \\ 1 & 2W_e x W_e y & W_e^2 y \\ 1 & 2W_f x W_f y & W_f^2 y \end{vmatrix} \\ &- 2 \begin{vmatrix} W_a^2 x & 1 & W_a^2 y \\ W_b^2 x & 1 & W_b^2 y \\ W_c^2 x & 1 & W_c^2 y \end{vmatrix} \begin{vmatrix} W_d^2 x & 1 & W_d^2 y \\ W_e^2 x & 1 & W_e^2 y \\ W_f^2 x & 1 & W_f^2 y \end{vmatrix}. \end{aligned} \tag{17}$$

In (17), $|\cdot|$ is the determinant, and the dependence of x and y on the parameter t has been omitted due to limitations of space. The function is a relative invariant, since it is proven in [4] to be a sum of products of the relative invariant functions $f_{i,j}(t)$ with $\{i, j\} \in \{a, b, c, d, e, f\}$. An absolute invariant is obtained in [4] by dividing the relative invariant $\eta(t)$ with another one composed of a different set of wavelet scales. Then, the total number of scales used ranges from 7 to 12 depending on how much overlap between the chosen scales of the two functions is allowed.

3.2. The affine-invariant multiwavelet function

With $MW_j x_i(t)$ denoting the multiwavelet coefficients (detail signal) of the x-coordinate function $x(t)$ of the preprocessed boundary at the ith row and scale j, and taking the multiwavelet transform of (11), it is observed that multiwavelet coefficients at identical rows, which correspond to the same wavelet $\psi_i(t)$, are related by an equation similar to the scalar wavelet case for the jth scale:

$$\begin{bmatrix} MW_j \tilde{x}_i(t) \\ MW_j \tilde{y}_i(t) \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} MW_j x_i(t) \\ MW_j y_i(t) \end{bmatrix}, \tag{18}$$

where $i \in \{1, 2, \ldots, R\}$. Thus, the affine-invariant function in (17) can be constructed from the multiwavelet coefficients by using six sets of coefficients of the form $MW_j x_i(t)$, where at each scale, there are R sets of coefficients.

3.3. The affine-invariant (multi)scaling function

The (multi)scaling function coefficients of the boundary curve depend on the position of the object in the image, since the effect of translation $\mathbf{b}$ in (12) is not eliminated with low-pass filtering. This dependence is, however, easily removed by selecting the centroid of the object as the center of the coordinate system when constructing the affine-invariant function. Then, the scaling and multiscaling function coefficients satisfy the same equations that the wavelet and multiwavelet coefficients do. Although the affine-invariant function $\eta(t)$ is constructed in a similar fashion from the (multi)wavelet and (multi)scaling function coefficients, there is a major difference between the invariants obtained. Whereas the (multi)wavelet coefficients at different scales correspond to orthogonal vector spaces W_j, the subspaces generated by the (multi)scaling functions are nonorthogonal. Therefore, the (multi)scaling function coefficients at different scales are related.

3.4. The choice of scales

In [4], where the affine-invariant function $\eta(t)$ is proposed, and constructed from wavelet coefficients, the first finest scales have been avoided because they are sensitive to noise. The effects of quantization are revealed in the finest scale of wavelet coefficients. The authors in [17], which advocates the use of scaling function coefficients instead of wavelet coefficients when constructing invariants, have observed that the amplitudes of the first few wavelet scales are small and highly sensitive to noise. As the scale gets coarser, the details of object boundaries have been removed from the (multi)scaling function coefficients by low-pass filtering. Thus, the more distinguishing features of objects are concentrated in the first finer-scale (multi)scaling function coefficients.

An object, which is known to belong to a specific database, is identified via the maximum normalized correlation value between its invariant function and the invariant functions of the objects in the database. Thus, six scales are enough to form the invariants, since they are matched by normalized correlation taking on values in the range $[0, 1]$.

The boundaries of objects are resampled to have a length of 2^7 and the redundant (multi)wavelet transform is taken for seven scales. For the multiwavelet transform, the RR preprocessing is employed with which better recognition performance is observed compared to AP preprocessing. This is a consequence of the fact that oversampled data representations are useful for feature extraction [15]. With RR preprocessing, the multiwavelet transform yields seven scales of coefficients for the object boundary like in the scalar case. The chosen scales for the experiments in the next section are as follows:

(i) wavelet: the finest scale is avoided and the coarsest six scales are chosen;
(ii) multiwavelet: both rows of the coarsest three scales are used with $R = 2$;
(iii) scaling function: the finest six scales are employed;
(iv) multiscaling function: the finest three scales with both rows of coefficients are used to construct the affine-invariant function with $R = 2$.

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, the recognition performance of the affine-invariant function $\eta(t)$ constructed with either (multi)wavelet or (multi)scaling function coefficients is investigated. Experiments using real airplane images, which are obtained from [4], are carried out with experimental setups similar to the ones in [4, 6]. The database shown in Figure 2 consists of 20 airplane images in their top view. It contains objects with very small differences, like models (g) and (t) or (r) and (s).

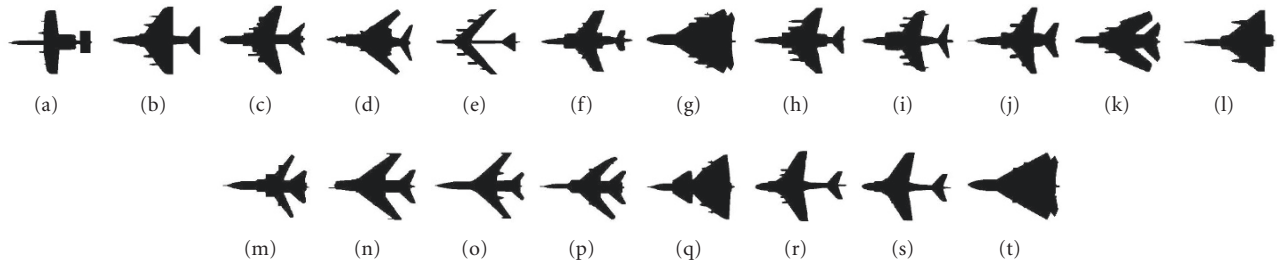

Figure 2: The database of airplane object images in their top view.

The boundaries of the objects are affine transformed with the transformation matrix

$$T = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 2 & b \\ 0 & \frac{1}{2} \end{bmatrix}, \tag{19}$$

where $\theta \in \{0°, 36°, \ldots, 324°\}$ and $b \in \{-2, -3/2, -1, \ldots, 3/2, 2, 5/2\}$ to realize the unknown object boundaries. This makes a total of 100 views for each reference object in the database. Initially, there exists perfect point correspondence between the boundary curves related by an affine transformation. For this ideal case of perfect point correspondence, the normalized correlation is exactly one. Before calculating the invariant function of the unknown object, uniformly distributed noise at an SNR of 20 dB is added to the boundary curve as in [4, 6]. SNR is defined as the ratio of the average squared distance of the boundary points from the centroid of the curve to the variance of the uniformly distributed noise. Uniform distribution takes on values from a finite range determined by its mean and variance. Adding uniformly distributed noise is realistic in the sense that any boundary tracking algorithm will find points closely spaced as belonging to the boundary. When the affine-transformed boundary of object (i) is disturbed with uniformly distributed white noise such that SNR = 20 dB, it appears as in Figure 3. The noise shifts the samples of the boundary in random directions. This results in a loss of point correspondence between the affine-transformed and reference curves, which lowers the correlation between their invariants.

Point correspondence between the two curves can be restored, for instance, by finding uniform starting points for the boundary curves of objects through registering the shapes based on an analysis of the discrete Fourier series phase differences [18], and subsequently employing an affine-invariant parametrization.

The experiments have been performed with four wavelets, two combined sets of wavelets and six multiwavelets:

(1) MZ: the wavelet in [7], which is used in [4];
(2) d4: Daubechies 4-coefficient orthogonal wavelet [19];
(3) la8: Daubechies 8-coefficient least asymmetric orthogonal wavelet [19];
(4) bi9: 9/7-coefficient symmetric biorthogonal wavelet [19];
(5) MZ-d4 combination;

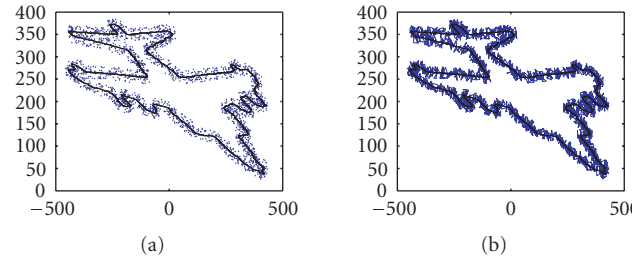

Figure 3: The noisy boundary of object (i) at SNR = 20 dB connected with (a) points and (b) lines.

(6) la8-bi9 combination;
(7) GHM: Geronimo-Hardin-Massopust orthogonal symmetric multiwavelet [20];
(8) CL: Chui-Lian orthogonal symmetric multiwavelet [21];
(9) SA4: orthogonal symmetric multiwavelet constructed by Shen et al. [22];
(10) bih52s: biorthogonal symmetric multiwavelet [23];
(11) bighm2: biorthogonal multiwavelet obtained from GHM by factoring out one approximation order [24];
(12) cardbal4: orthogonal cardinal 4-balanced multiwavelet constructed by Selesnick [25].

For the MZ-d4 and la8-bi9 combinations, the wavelet-coefficient based-invariant is calculated from the coarsest three scales and when the scaling function coefficients are available, the finest three scales are made use of. Coefficients of two wavelets are combined in one invariant in an effort to make the number of single wavelets applied equal to those of multiwavelets: the multiwavelets above have $R = 2$. The equations that multiwavelets have to satisfy make it difficult to construct multiwavelets with $R > 2$.

The boundaries of the objects in the database are affine transformed using (19), and different realizations of noise are added to each transformed curve such that SNR = 20 dB. The affine-invariant function $\eta(t)$ calculated from the transformed curve, which is noisy, is correlated with the invariants of the objects in Figure 2 via

$$\frac{\sum_{t=0}^{N-1} \eta(t)\eta_i(t)}{\sqrt{\sum_{t=0}^{N-1} \eta^2(t) \sum_{t=0}^{N-1} \eta_i^2(t)}}, \tag{20}$$

Table 1: Number of correctly matched poses of objects at 20 dB with (multi)wavelet coefficients.

	Number of correct matches											
Plane	MZ	d4	la8	bi9	MZ-d4	la8-bi9	GHM	CL	SA4	bih52s	bighm2	cardbal4
(a)	100	87	74	90	100	99	100	100	97	100	100	91
(b)	100	64	71	60	99	78	99	99	100	91	93	99
(c)	100	69	67	62	96	84	98	100	95	95	95	79
(d)	88	72	45	57	94	83	97	98	96	96	87	48
(e)	100	84	84	63	100	95	100	100	100	100	100	100
(f)	94	74	78	62	100	89	100	100	95	98	96	90
(g)	87	42	47	32	99	93	93	95	93	89	82	79
(h)	99	89	73	61	100	100	99	99	98	93	94	95
(i)	97	73	75	72	97	74	88	98	97	93	95	79
(j)	95	68	62	45	100	94	97	93	92	94	95	88
(k)	93	65	61	67	98	85	94	96	93	93	90	75
(l)	84	87	65	44	97	71	99	96	96	96	90	84
(m)	93	57	55	39	98	88	96	95	95	95	90	91
(n)	94	67	66	62	99	84	90	99	100	93	96	70
(o)	97	75	57	48	98	95	99	99	100	99	99	94
(p)	84	55	52	34	98	81	89	94	92	92	81	74
(q)	82	57	69	49	100	76	100	100	99	98	88	80
(r)	100	76	76	63	100	85	97	99	100	100	100	96
(s)	98	69	56	43	99	84	100	99	99	100	97	84
(t)	73	48	16	14	100	74	94	89	99	81	82	47
Average	92.9	68.9	62.5	53.4	98.6	85.6	96.5	97.4	96.8	94.8	92.5	82.15

where $i \in \{1, \ldots, 20\}$, $\eta(t)$ and $\eta_i(t)$ of length 2^7 are the invariants of the unknown and reference objects, respectively. The maximum of the correlations identifies the unknown object.

The recognition performance of the affine-invariant function constructed from (multi)wavelet coefficients is displayed in Table 1 as the number of correct matches for each object. In the last row, the average (multi)wavelet performance is shown. Although the highest average recognition rate is achieved by MZ-d4 combination, the multiwavelets have generally outperformed the scalar wavelets. The function based on a combination of la8 and bi9 coefficients is a major improvement over the la8 and bi9 invariants. In addition, the recognition rates of the (multi)wavelets are seen to be correlated with the objects, where, for instance, plane (t) generally has the lowest rates among all of the objects. These observations are related to the approximation properties of (multi)wavelets.

The subspaces V_j spanned by translates of (multi)scaling functions are required to reproduce polynomials up to a certain degree $K - 1$ [26]. Therefore, as W_j is orthogonal to V_j, the first K moments of the (multi)wavelet vanish,

$$\int t^k \Psi(t)dt = 0, \quad k = 0, \ldots, K-1. \tag{21}$$

Such a (multi)wavelet has approximation order K. When $R = 1$, the span of $\phi(t)$ contains all polynomials of degree $< K$. However, when $R > 1$, the span of each individual scaling function $\phi_i(t)$, $i \in \{1, \ldots, R\}$, does not have to contain all such polynomials [26]. For $R = 1$, the degree of polynomials that can be exactly represented by a sum of weighted and shifted scaling functions is shown to be tied to the number of zero moments of wavelet filters as well [14]: all moments of the wavelet (high-pass) filters are zero, $\mu(k) = 0$ for $k = 0, 1, \ldots, K-1$, where

$$\mu(k) = \sum_n n^k g(n). \tag{22}$$

(Multi)wavelets in this work have approximation orders of 1, 2, or 4:

$K = 1$: MZ, SA4, bighm2;
$K = 2$: d4, GHM, CL, bih52s;
$K = 4$: la8, bi9, cardbal4.

A higher approximation order necessitates an increase in the length of filter coefficients for wavelets as exemplified by d4 and la8 wavelets.

Theoretically, smoother objects can be represented by polynomials of lower degree. The (multi)scaling function coefficients of such objects are sufficient for an accurate representation and the (multi)wavelet coefficients, which show the details, have small amplitudes. Hence, the affine-invariant function $\eta(t)$ using (multi)wavelet coefficients of the (multi)wavelets with higher K fails to recognize smoother objects at a corresponding higher rate. MZ wavelet with the lowest K is the most successful wavelet in terms of recognition performance. d4 with $K = 2$ comes next and the two other wavelets, la8 and bi9 having $K = 4$, are especially unsuccessful with the smoothest object, plane (t). Combining

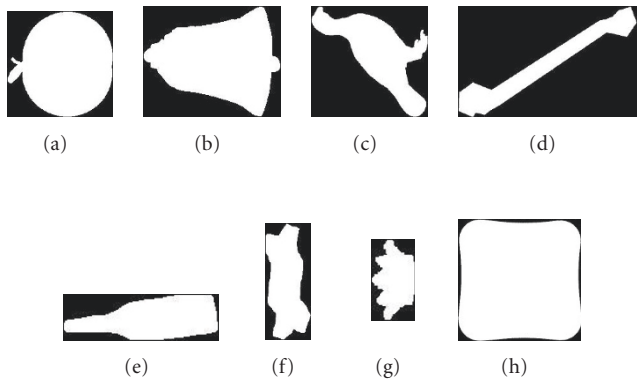

Figure 4: The database of smooth object images.

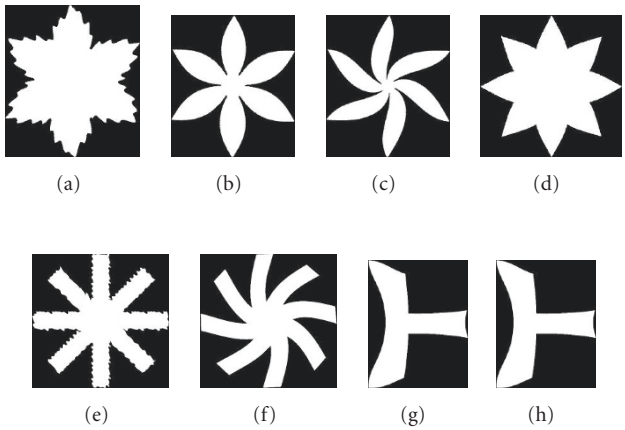

Figure 5: The database of device images.

their wavelet coefficients at the coarser scales so that the finest scales with small amplitudes are avoided helps to improve their joint performance when using the wavelet-coefficient based invariant function. The multiwavelets have generally high recognition rates, since their increased dimensionality makes it possible to disregard the noisy finest scales in the beginning. The lower number of correct matches for cardbal4 with $K = 4$ justifies the claim about the approximation properties of (multi)wavelets.

Two image databases have been formed to separate smooth objects from those that contain more details. They are shown in Figures 4 and 5, respectively. The same experiment is repeated and the average recognition performance of (multi)wavelets is given in Table 2. As expected, the average number of correct matches for wavelets with high K (i.e., la8 and bi9) is very low for smooth objects. The objects in the device images database have been correctly matched a larger number of times, since their details have been captured by the (multi)wavelet coefficients.

Employment of (multi)scaling coefficients when constructing the affine-invariant $\eta(t)$ affects its performance so much that wavelets with high K should be preferred over the multiwavelets. The average recognition rates for the three databases are given in Table 3. For smooth objects, combining the scaling function coefficients of two wavelets deteriorates the performance of the invariant while this observation does not hold for the device objects. Among the multiwavelets, cardbal4 has exceptionally good recognition performance. The average number of correct matches with GHM is high as well.

As opposed to (multi)wavelet coefficients, which are coefficients of basis functions spanning orthogonal spaces, (multi)scaling function coefficients are obtained from nested subspaces V_j which are related. Therefore, when the coefficients of two different scaling functions are combined, as in each of the MZ-d4, la8-bi9, or multiwavelet cases, it is essential that the scaling function filters and the approximation properties of the individual scaling functions are similar. Specifically, a constant signal should remain constant after filtering. Filtering a constant signal $c = 1$ with the lowpass filters of the wavelets and multiwavelets used here produces the signals c_ℓ and $[c_{\ell,1} \;\; c_{\ell,2}]^T$, respectively, where T is the transpose:

(i) $c_\ell = 1$: MZ;
(ii) $c_\ell = \sqrt{2}$: d4, la8, bi9;
(iii) $c_{\ell,1} = \sqrt{2}c_{\ell,2} = 1$: GHM;
(iv) $c_{\ell,1} = \sqrt{2}c_{\ell,2} = 0$: SA4, CL, bih52s, bighm2;
(v) $c_{\ell,1} = \sqrt{2}c_{\ell,2} = \sqrt{2}$: cardbal4.

Hence, for some of the multiwavelets, the multiscaling function coefficients of the boundaries of smoother objects at different rows should not be jointly used in calculating the affine-invariant function $\eta(t)$ which requires them to be multiplied as in (17). The result is especially catastrophic for the SA4, CL, bih52s, and bighm2 multiwavelets. The average number of correct matches in Table 3 is in compliance with this observation and cardbal4 rightfully achieves high recognition rates with an accompanying good performance for GHM.

The performance of the affine-invariant $\eta(t)$ is a function of the amount of noise on the boundary curve of the object as well. Thus, a final experiment is made, where the realizations of noise are produced in such a way that SNR is varied between 20 dB and 50 dB. The results are shown in Figure 6. The averages of the correlations (20) between the invariants obtained from the reference and noisy-affine transformed curves for the objects in Figure 2 are displayed. It is observed that the invariants based on either the scaling function or the multiwavelet and the combined sets of wavelet coefficients are less sensitive to the amount of noise.

5. CONCLUSION

In this work, recognition of planar objects from their test images which have been obtained from different directions than their standard view in a database has been considered. The test images and the reference ones in the database are related by an affine transformation. Thus, affine invariants are required for recognition. Previously, an affine-invariant function calculated from the wavelet coefficients of the object boundary has been proposed in [4]. However,

Table 2: The average number of correct matches at 20 dB with (multi)wavelet coefficients.

Object	MZ	d4	la8	bi9	MZ-d4	la8-bi9	GHM	SA4	CL	bih52s	bighm2	cardbal4
Smooth	94.3	46.4	37.4	33.5	99.1	92.5	94.9	81.5	93.4	80.8	89.4	72.9
Device	100.0	90.4	84.5	72.0	91.6	79.3	96.4	90.8	95.8	91.4	96.1	96.0

Table 3: The average number of correct matches at 20 dB with (multi)scaling function coefficients.

Object	MZ	d4	la8	bi9	MZ-d4	la8-bi9	GHM	SA4	CL	bih52s	bighm2	cardbal4
Plane	97.2	96.9	94.3	95.4	78.9	78.0	65.9	29.6	36.9	13.4	14.5	90.0
Smooth	93.1	99.5	99.5	89.6	80.3	80.5	80.4	38.0	41.1	26.8	21.6	93.6
Device	98.5	100.0	98.4	66.5	93.6	93.3	95.0	60.8	63.4	54.9	39.0	97.1

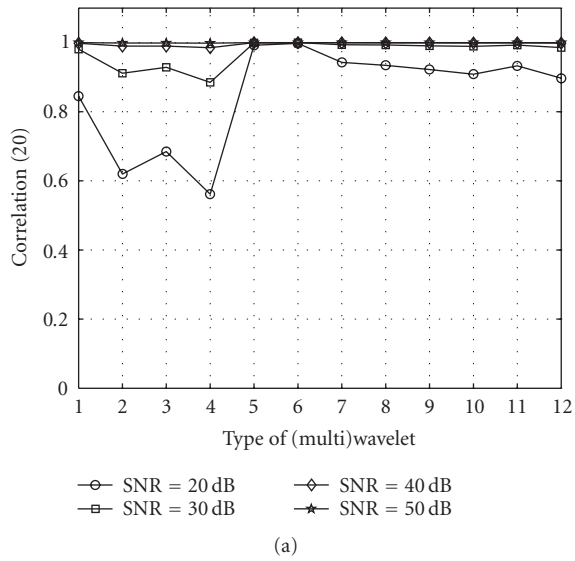

Figure 6: The average correlation values (20) for the plane objects with (a) multiwavelet coefficients and (b) multiscaling function coefficients; x-axis corresponds to the (multi)wavelets used in this section.

the performance of the function has been tested for only one view of the object with one type of wavelet. Here, we have calculated the invariant from (multi)wavelet and (multi)scaling function coefficients of the boundary. An extensive set of simulations are made, which indicate the following:

(i) multiwavelets have a superior performance compared to scalar wavelets when the detail coefficients are available. For smooth objects, the result is more pronounced;

(ii) the scaling function coefficients of two different scaling functions should not be used jointly in one invariant function due to the fact that the scaling function coefficients at different scales are expected to be correlated;

(iii) the scaling function coefficients are more stable compared to wavelet coefficients in the presence of noise.

The observations above are shown to be closely related to multiresolution analysis and the approximation properties of (multi)wavelets.

ACKNOWLEDGMENT

The authors would like to thank Vasily Strela for generously providing his multiwavelet software package (MWMP), where the coefficients for most of the (multi)wavelets used in this paper can be found.

REFERENCES

[1] I. Weiss, "Geometric invariants and object recognition," *International Journal of Computer Vision*, vol. 10, no. 3, pp. 207–231, 1993.

[2] K. Arbter, W. E. Snyder, H. Burkhardt, and G. Hirzinger, "Application of affine-invariant Fourier descriptors to recognition of 3-D objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 640–647, 1990.

[3] R. Alferez and Y.-F. Wang, "Geometric and illumination invariants for object recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 6, pp. 505–536, 1999.

[4] M. I. Khalil and M. M. Bayoumi, "A dyadic wavelet affine invariant function for 2D shape recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1152–1164, 2001.

[5] Q. M. Tieng and W. W. Boles, "Wavelet-based affine invariant representation: a tool for recognizing planar objects in 3D space," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 8, pp. 846–857, 1997.

[6] E. Bala and A. E. Cetin, "Computationally efficient wavelet affine invariant functions for shape recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 8, pp. 1095–1099, 2004.

[7] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 710–732, 1992.

[8] M. B. Martin and A. E. Bell, "New image compression techniques using multiwavelets and multiwavelet packets," *IEEE Transactions on Image Processing*, vol. 10, no. 4, pp. 500–510, 2001.

[9] T. D. Bui and G. Chen, "Translation-invariant denoising using multiwavelets," *IEEE Transactions on Signal Processing*, vol. 46, no. 12, pp. 3414–3420, 1998.

[10] E. Bala and A. Ertüzün, "A multivariate thresholding technique for image denoising using multiwavelets," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 8, pp. 1205–1211, 2005.

[11] F. P. Nava and A. F. Martel, "Planar shape representation based on multiwavelets," in *Proceedings of 10th European Signal Processing Conference (EUSIPCO '00)*, Tampere, Finland, September 2000.

[12] T. N. T. Goodman and S. L. Lee, "Wavelets of multiplicity r," *Transactions of the American Mathematical Society*, vol. 342, no. 1, pp. 307–324, 1994.

[13] V. Strela, P. N. Heller, G. Strang, P. Topiwala, and C. Heil, "The application of multiwavelet filterbanks to image processing," *IEEE Transactions on Image Processing*, vol. 8, no. 4, pp. 548–563, 1999.

[14] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms*, Prentice-Hall, Upper Saddle River, NJ, USA, 1998.

[15] V. Strela, *Multiwavelets: theory and applications*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Mass, USA, 1996.

[16] S. Mallat, "Zero-crossings of a wavelet transform," *IEEE Transactions on Information Theory*, vol. 37, no. 4, pp. 1019–1033, 1991.

[17] I. El Rube, M. Ahmed, and M. Kamel, "Wavelet approximation-based affine invariant shape representation functions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 323–327, 2006.

[18] M. J. Paulik and Y. D. Wang, "Three-dimensional object recognition using vector wavelets," in *Proceedings of the IEEE International Conference on Image Processing*, vol. 3, pp. 586–590, Chicago, Ill, USA, October 1998.

[19] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, Pa, USA, 1992.

[20] G. S. Geronimo, D. P. Hardin, and P. R. Massopust, "Fractal functions and wavelet expansions based on several functions," *Journal of Approximation Theory*, vol. 78, no. 3, pp. 373–401, 1994.

[21] C. K. Chui and J. A. Lian, "A study of orthonormal multiwavelets," CAT Report 351, Texas A&M University, Canyon, Tex, USA, 1995.

[22] L.-X. Shen, H. H. Tan, and J. Y. Tham, "Symmetric-antisymmetric orthonormal multiwavelets and related scalar wavelets," *Applied and Computational Harmonic Analysis*, vol. 8, no. 3, pp. 258–279, 2000.

[23] R. Turcajova and V. Strela, "Smooth hermite spline multiwavelets," in preparation.

[24] V. Strela, "A note on construction of biorthogonal multiscaling functions," in *Contemporary Mathematics*, A. Aldroubi and E. B. Lin, Eds., vol. 216, pp. 149–157, American Mathematical Society, Providence, RI, USA, 1998.

[25] I. Selesnick, "Cardinal multiwavelets and the sampling theorem," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '99)*, vol. 3, pp. 1209–1212, Phoenix, Ariz, USA, March 1999.

[26] K. Berkner and P. R. Massopust, "Translation invariant multiwavelet transforms," Tech. Rep. CML TR 98-06, Computational Mathematics Laboratory, Rice University, Houston, Tex, USA, 1998.

Nazlı Güney received the B.S. (with high honors) and M.S. degrees in electrical and electronics engineering from Boğaziçi University, Istanbul, Turkey, in 2001 and 2003, respectively, and she is currently working toward the Ph.D. degree in electrical and electronics engineering from Boğaziçi University. Since 2001, she has been a Research and Teaching Assistant with Boğaziçi University. She worked on planar object recognition for her M.S. thesis. Her current research interests include various aspects of UWB communications with special emphasis on design and analysis of robust systems for non-Gaussian channels.

Ayşın Ertüzün was born in 1959 in Salihli, Turkey. She received the B.S. degree (with honors) from Boğaziçi University, Istanbul, Turkey, the M.Eng. degree from McMaster University, Hamilton, Ontario, Canada, and the Ph.D. degree from Boğaziçi University, Istanbul, Turkey, all in electrical engineering, in 1981, in 1984, and in 1989, respectively. Since 1988, she has been with the Department of Electrical and Electronics Engineering at Boğaziçi University where she is currently a Professor. Her current research interests are in the areas of independent component analysis and its applications, blind signal processing, Bayesian methods, application of wavelets and adaptive systems to communication systems, image processing, and texture analysis. She has authored and coauthored nearly 70 scientific papers in journals and conference proceedings. She is a Member of IEEE Signal Processing and Communication Societies, International Association of Pattern Recognition (IAPR), The Institute of Electronics, Information and Communication Engineers (IEICE), and Turkish Pattern Recognition and Image Processing Society (TOTIAD).